DOCUMENT RESUME

ED 439 548 EC 307 727

AUTHOR	Cullen, Joe; Shaw, Stan
TITLE	The Accuracy of Teacher Prediction of Student Test Performance for Students Referred to Special Education.
PUB DATE	2000-00-00
NOTE	41p.
PUB TYPE	Reports - Research (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Diagnostic Tests; *Disabilities; *Disability Identification; Elementary Secondary Education; *Eligibility; Predictor Variables; Psychological Testing; *Referral; Special Education; Student Placement; Teacher Attitudes; *Teacher Expectations of Students; Teacher Surveys
IDENTIFIERS	Connecticut

ABSTRACT

A survey was conducted of 55 regular classroom teachers who referred students to special education in three Connecticut school systems to determine their predictions of the results of psycho-educational assessment and their perceptions of the instructional relevance of standardized psychological and educational test results. The findings of this study indicate that teachers could accurately predict the standardized test performance of students referred for special education using the same standard score and percentile rank formats employed by school psychologists and other diagnosticians. It is, therefore, argued that psychological and educational reports that focus primarily or exclusively on test scores may do little more than confirm teachers' impressions of their students. This investigation also found that teachers are ambivalent about the instructional value of psycho-educational assessment. Overall, these results are used to argue for greater teacher input into special education identification and placement procedures. These would include more in-depth involvement of specialists in the pre-referral stage of the eligibility process and a stronger focus on assessment domains that are more likely to yield new and useful information for teachers, such as affective development, behavioral adjustment, motivation, information processing, and data-based intervention strategies. Data tables are attached. (Contains 47 references.) (Author/CR)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 439 548

THE ACCURACY OF TEACHER PREDICTION OF STUDENT TEST PERFORMANCE FOR STUDENTS REFERRED TO SPECIAL EDUCATION

Joe Cullen, Ph.D.

Department of Education & Educational Psychology

Western Connecticut State University

Stan Shaw, Ed.D.

Department of Educational Psychology

The University of Connecticut

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Cullen

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Running head: TEACHER PREDICTION

EC 307727

BEST COPY AVAILABLE

Abstract

A survey was conducted of 55 regular classroom teachers who referred students to special education in three Connecticut school systems to determine (a) their predictions of the results of psycho-educational assessment and (b) their perceptions of the instructional relevance of standardized psychological and educational test results. The findings of this study indicate that teachers can accurately predict the standardized test performance of students referred for special education using the same standard score and percentile rank formats employed by school psychologists and other diagnosticians. It is, therefore, argued that psychological and educational reports that focus primarily or exclusively on test scores may do little more that confirm teachers' *a priori* impressions of their students. This investigation also found that teachers are ambivalent about the instructional value of psycho-educational assessment. Overall, these results are used to argue for greater teacher input into special education identification and placement procedures, more in-depth involvement of specialists in the pre-referral stage of the eligibility process, and a stronger focus on assessment domains that are more likely to yield "new" and useful information for teachers; e.g., affective development, behavioral adjustment, motivation, information processing, and data-based intervention strategies.

The Accuracy of Teacher Prediction of Student Test Performance for Students Referred to Special Education

Since its inception, the field of special education has struggled with issues of eligibility for services and adequacy of diagnostic assessment. In recent years, these topics have emerged in the form of concerns about over-identification, particularly in the area of learning disabilities, and the instructional value of psychological and educational testing. Because of its bearing on the quality of initial referral decisions and the value of diagnostic test results, a teacher's ability to judge students' cognitive and academic functioning is thought to have important implications for these concerns. Therefore, this study will explore the accuracy of teacher judgment of students' abilities and its meaning for referral and placement practices.

The decision whether or not to refer a student for special education has been consistently found to be a powerful predictor of subsequent special education placement (Thurlow, Christenson, & Ysseldyke, 1983; Ysseldyke & Algozzine, 1981; Ysseldyke, Algozzine, Regan, & McGue, 1981). Therefore, the quality of identification and placement procedures in special education may be strongly influenced by the quality of initial referral decisions. In addition, if it is assumed that most special education referrals are the result of concerns about student abilities

in one domain or another, it follows that the accuracy of teacher judgments of student abilities may play a key role in determining the quality of initial referral decisions. Over the past 20 years, a number of studies have examined the accuracy of teacher judgments of student abilities, including the accuracy with which they estimate academic and cognitive functioning and the credibility of their judgments of handicapping conditions. Unfortunately, the majority of these investigations have focused on generic student samples rather than samples of students referred for special education services.

Estimating students' academic and cognitive functioning

While they differ markedly in approach and format, investigations of teachers' ability to estimate or predict the academic or cognitive functioning of their students have primarily used correlational methods. Early efforts by Dusek and O'Connell (1973) and Morine-Dershimer (1978-79) used a predictive approach that compared teacher groupings with actual end-of-year test scores. They found that there was a strong relationship between teacher expectations for student achievement at the beginning of the school year and actual performance on an achievement measure administered at year's end.

Others employed a concurrent format that compared either categorical or rank order estimates of student abilities with either the numerical results of standardized

tests or students' responses to specific test items (Coladarci, 1986; Kellaghan, Madaus, & Airasian, 1982; Luce & Hoge, 1978; Mayfield, 1979; Miller & Davis, 1992; Mulholland & Berliner, 1992; Oliver & Arnold, 1978; Sharpley & Edgar, 1986; Wilson & Wright, 1993). Hoge and Butcher (1984), in a review of such studies from the mid 70's through early 80's, found that they reported a median correlation of .55 between teacher estimates of student ability and actual performance on standardized tests.

This analysis of the literature is consistent with investigations that have been conducted since that time. It should, however, be noted that in a number of these studies, teachers were significantly less accurate in estimating the achievement and ability of lower functioning students. Several of these investigations also found that experienced teachers provided significantly more accurate predictions than their novice counterparts (Mulholland and Berliner, 1992; Stader, 1990).

Another significant source of research on the accuracy of teacher predictions of student performance on standardized measures of achievement and intelligence are investigations attempting to validate the tests themselves. Webster, Hewett, and Crumbacker (1989) found moderate correlation coefficients in the .4 to .6 range between teacher estimates of academic achievement and student performance on the Wide Range Achievement Test (Jastak & Wilkinson, 1985) and Kaufman Test of Educational Achievement

(Kaufman & Kaufman, 1985). Similar results were obtained by Hopkins, George, and Williams (1985) in a study of the concurrent validity of the Comprehensive Test of Basic Skills (Comprehensive Test of Basic Skills, 1984). With regard to tests of mental ability, Reilly, Drudge, Rosen, Loew, and Fischer (1985) demonstrated the concurrent validity of the Wechsler Intelligence Test for Children - Revised (Wechsler, 1974), McCarthy Scale of Children's Abilities (McCarthy, 1972), and Woodcock-Johnson Psycho-educational Battery (Woodcock & Johnson, 1977) by correlating them with teacher estimates of children's intellectual functioning.

Judging students as handicapped

Another approach to evaluating the accuracy of teacher judgments of student abilities was exemplified by Gerber and Semmel (1984) who examined teacher ability to judge students' handicapping conditions. Their finding that teachers were not accurate judges of the existence of learning handicaps was consistent with the results of a study by Rothlisberg and Liljestrom (1984) which determined that a majority of teachers held misconceptions about the term "learning disability." These two investigations questioned educators' ability and purpose in making classification and placement decisions. On the other hand, a study by Valus (1986), reported that teachers had very accurate perceptions of identification criteria used in

initial learning disabilities placements. Nonetheless, the preponderance of the evidence suggests that teachers are not as accurate in judging students handicapping conditions as they are in estimating their cognitive and academic functioning.

Implications for diagnostic assessment practices

Generally speaking, the literature on teacher judgment of students' standardized test performance provides strong evidence for the validity of teacher's concurrent and predictive estimates. The strength of this research has prompted some to conclude that teacher judgments of student academic performance are as accurate and useful as standardized test results in making classification, placement, and programming decisions (Gresham, Reschly, & Carey, 1987; Quay, 1996; Schendel & Binder-Reschly, 1989). It is suggested that a teacher who spends six hours a day with a student for a period of months has insights into the child's functioning that cannot be gleaned from two hours of psychometric testing. As Gresham et al. (1987) write, "Teacher judgments most certainly are based upon a much wider and more comprehensive sampling of the content domain of achievement and classroom behavior than standardized tests" (p. 544).

Implications for the relevance of psychological and educational testing

Given that teachers have been found to be accurate predictors of student performance on psycho-educational tests, the question then becomes "How valuable do they find psycho-educational reports that dwell primarily on information they already know?" Unfortunately, the literature on teacher perceptions of psycho-educational reports is rather limited. Ysseldyke and Shinn (1981) reported the results of a 1968 survey in which nearly 6,000 teachers in New York were asked to rate the helpfulness of school psychology reports. Approximately 77% rated them as either "of no help," "not relevant," or "detrimental."

While there have been enormous changes in special education and the profession of school psychology since the time of this investigation, more current research indicates that teachers are, at best, "tepid" in their opinions of diagnostic evaluations. Of particular concern is the observation that psychological and educational reports focus primarily on standardized test scores rather than material that is more instructionally relevant; such as clear, jargon-free statements about the nature of a child's difficulties and practical intervention approaches (Noble & Dickinson, 1988; Ownby, Wallbrown, & Brown, 1982; Weiner, 1985, 1987). Zins and Barnett (1983), in a review of the literature on psychological report writing, wrote:

> Often we find that the results are not helpful in decision-making, intervention strategies do not follow from the report, teachers and administrators do not read them, and special educators complain that they cannot translate the information directly into the IEPs. (p. 129)

In summary, this investigation makes two assumptions that are supported by the literature: (1) the quality of identification and placement practices in special education is heavily affected by the quality of initial referral decisions made by teachers and (2) a factor that significantly influences the quality of initial referral decisions is the accuracy of teacher judgments of students abilities. Therefore, research aimed at improving special education referral and placement practices should examine this issue. Prior research on this subject has focused on generic samples and produced mixed results. However, it generally supports the conclusion that teachers are accurate judges of students' cognitive and academic abilities though they are less skilled at correctly identifying handicapping conditions. The literature has also found that teachers often do not perceive psycho-educational assessment to be of much instructional value and that they would benefit from a greater focus on prescription; that is, the delineation of appropriate intervention strategies and objectives.

To extend the literature on teacher judgment of student abilities, this investigation will examine the relationship between teacher estimates of the academic and cognitive functioning of students referred for special education

services and the subsequent performance of those students on standardized measures of achievement and intelligence. While some investigations have found teacher estimates to be less accurate on lower functioning students and the majority of students referred for special education are, in fact, low functioning students, this study will attempt to answer the question "Does the accuracy of teacher estimates of the cognitive and academic abilities of students referred for special education match the standards found in the literature for other student groups?"

In addition, since most of the literature on teacher prediction uses ranking or categorical rating procedures, very little is known about the ability of teachers to estimate student performance in the form most commonly used by diagnosticians--standardized scores and percentile ranks. This study will, therefore, ask "Can teachers accurately estimate student ability and achievement using percentile ranks and standard scores with a mean of 100 and a standard deviation of 15?"

Finally, this study includes a procedure for evaluating teacher perceptions of the educational value of intelligence and achievement test data. This information will be used to address the question "To what extent do teachers find intelligence and achievement testing instructionally useful versus institutionally necessary?" The results of this study will then be discussed with respect to the special education identification and placement process.

Method

Instrumentation

To evaluate teacher accuracy in estimating student intellectual and academic skills, a pool of potential survey items was developed and reviewed by experts in the field. Based on their feedback, a 19-item draft instrument was developed and piloted. The pilot data was then used to develop a final instrument which included nine items asking for demographic information, five dealing with teacher estimates of students' intellectual and academic functioning, and three that evaluated teacher attitudes toward psycho-educational assessment.

Demographic questions were based on a multiple choice format with one fill-in-the-blank "other" option to assure that all possible responses would be covered. They included items that asked for the respondents' years of experience, type of teaching certificate, current degree status, pre-referral interventions attempted, reason for referral, and whether or not previous intelligence and/or achievement tests results had been reviewed.

The format for items that asked teachers to rate students' abilities was mixed. It included a multiple choice estimate of student aptitude in which respondents checked one of seven categories ranging from "very poor" (IQ under 70) to "very superior" (IQ of 130 or higher); a fill-in-the-blank estimate of student aptitude in which teachers

were asked to supply a predicted score for the child on an intelligence test with a mean of 100 and a standard deviation of 15; estimates of student achievement in Mathematics, Reading, and Written Language arranged in eight categories extending from "below the second percentile" to "98th percentile and up;" and estimates of student achievement in which respondents were asked to enter a grade level score in 11 separate content areas (math computation, math concepts, math overall, reading sight-word, reading decoding, reading comprehension, reading overall, written language grammar, written language spelling, written language expression, and written language overall).

To address the issue of how teachers perceive intelligence and achievement testing, the examiners asked respondents to indicate the extent to which they agreed with the following three statements:

(1) Information from Intelligence/Achievement testing improves my understanding of the students who are assessed.

(2) Information from Intelligence/Achievement testing helps me to teach students more effectively.

(3) The primary purpose of Intelligence/Achievement testing is to determine whether or not a student qualifies for special education.

Agreement was indicated on a nine-point Likert scale ranging from Strongly Disagree (1) to Strongly Agree (9). A sample copy of the survey is contain in Figure 1.

Insert Figure 1 about here

Participants

In all, 68 surveys were distributed to five elementary schools in three eastern Connecticut school systems; one urban, one rural, and one rural/suburban. The three school systems participating in this study serve communities with a combined population of approximately 60,000, 13.25% of whom live below the poverty level. Approximately 11% of the adults over age 25 in these communities have college degrees while 61% hold high school diplomas. Because approximately 91% of the families in the three communities send their children to public school and a majority of the elementary schools in the three school systems participated in the study, the sample used for this investigation is thought to be highly representative of the communities from which it was drawn.

The three school systems that participated in the study report a 10.6% minority and 2.5% non-English speaking enrollment. Approximately 15% of their students are described as economically disadvantaged and 13.6% have been found eligible for special education and related services.

The average per pupil expenditure during the 1993-94 school year (the most recent year for which data were available) was $6588 per year. In terms of achievement levels, approximately 12% of the students in the participating schools scored below the state's remedial standard in math, 22% in reading, and 13% in writing. These statistics place them well within state norms.

Procedure

Surveys were given to teachers who had referred students for special education placement but had not yet reviewed the results of psycho-educational assessments. The participants were instructed to make specific estimates of student functioning based on their professional judgment and experience, including their observations and samples of student work. Respondents were assured of anonymity and confidentiality. Given that any subsequent inquiries would have been biased by teachers exposure to the results of psycho-educational assessment, follow-up procedures were considered inappropriate.

Data Analyses

Frequency distributions and descriptive statistics were derived for the demographic and attitudinal items. This data was used to profile the sample and distribution of practitioner responses. To quantify the accuracy of teacher

estimates of student performance on standardized tests, Pearson Product-Moment correlation coefficients were computed. Differences between different groups of respondents with respect to the dependent variables were explored via analysis of variance (ANOVA) and t-test procedures. Because the sample for this investigation was small, a .05 level of significance was used for all statistical tests. Finally, to provide a more practical/meaningful measure of the efficacy of teacher predictions of student performance on standardized intelligence tests, this investigation will report a "hit rate" that reflects the degree of correspondence between eligibility data and teacher predictions.

Results

Response rate

A total of 58 out of 68 surveys were returned for an 85% response rate. Of the those returned, 55 were matched with psycho-educational report data to run the correlation, t-test, and ANOVA procedures discussed below. However, due to missing data, some teacher estimates were not reportable (due to small N's) and others were based on less than the full sample of 55 complete cases.

Sample

The sample of respondents for this investigation was composed primarily of highly experienced and trained educators. In terms of teaching experience, the vast majority of respondents (78.2%) indicated that they had taught for more than 10 years. With regard to educational background, nearly 73% of respondents held advanced degrees (masters, sixth year certificates, or doctorates). In addition, more than 83% of the respondents held standard (i.e., permanent) teaching certificates while approximately 17% held provisional (i.e., initial) licensure.

Descriptive Statistics

In response to an item asking if teachers had reviewed the results of previously administered intelligence or achievement measures, six of 55 subjects indicated that they had done so. In most instances, these responses referred to group-administered tests such as the Metropolitan Achievement Test (Prescott, Barlow, Hogan, & Farr, 1987) or Comprehensive Test of Basic Skills (CTBS 1984, 1990). Nonetheless, a check on the possibility of bias resulting from this factor will be explored under the analysis of variance/t-test results section.

Teacher responses to an item asking them to indicate the extent of their contact with students prior to referral indicated a mean of approximately five months. In addition, data on the frequency with which interventions were attempted prior to referral indicated that extensive pre-

referral work was done in most instances; e.g., 86% of respondents had reviewed cumulative records, 95% had done classroom observations, 89% had analyzed student work samples, and 78% had done informal (i.e., curriculum-based or criterion-referenced) assessment. These findings suggest that respondents had an extensive knowledge base from which to draw in making their estimates of students' intellectual and academic functioning.

The 55 student referrals analyzed for this investigation ranged in age from 5 years, 2 months through to 14 years, 11 months with a mean of 9 years, 1 month. With respect to grade, the sample ranged from preschool through grade seven with modes at the first and fourth grade levels. Roughly 60% of the cases were from the primary grades with the remaining 40% distributed between the elementary and middle school levels. Approximately two-thirds of the referrals were made in the Fall of the year (i.e., from September through December) with the remaining third coming in the Winter and Spring (i.e., January through June). The gender ratio for the sample was approximately 2.3:1 in favor of males.

The results of psychometric assessments used as criterion variables indicate that the sample of student referrals in this investigation performed on standardized intelligence tests within what is traditionally considered to be the "average" range. The mean full-scale intelligence quotient for the sample was approximately 93 with a range of

68 to 125 (mean of 100, standard deviation of 15). On standardized achievement tests, the mode for academic performance in mathematics, reading, and written language was the "low average" range; that is, percentile rank scores ranging from 9 to 25 and standard scores between 80 and 89 on a scale with a mean of 100 and a standard deviation of 15.

With regard to the question "To what extent do teachers find intelligence and achievement testing instructionally useful versus institutionally necessary?", results indicate that the teachers sampled may be ambivalent about the instructional value of psycho-educational assessment. More specifically, approximately 60% of the respondents agreed or strongly agreed with the statement that "information from intelligence/achievement testing improves my understanding of the students who are assessed" and 51% agreed or strongly agreed with the statement that "information from intelligence/achievement testing helps me to teach students more effectively." However, approximately 69% of the respondents indicated that they either strongly agreed with, agreed with, or had no opinion regarding the statement "the primary purpose of intelligence/achievement testing is to determine whether or not a student qualifies for special education." In addition, the mean scores on the three attitudinal items of 6.3, 6.0, and 5.3, respectively, all fell within the "uncertain" range on the 9-point scale utilized in the survey. These findings suggest that, while

they perceive psycho-educational assessment to be of some instructional value, the teachers sampled were as likely to view it as an institutional necessity (i.e., the determination of eligibility) as they were to see it as a meaningful prescriptive endeavor.

Insert Table 1 about here

Correlation

Concerning the issue of what teachers know at the time of referral, data on teacher estimates of student ability and achievement were matched with actual scores from standardized intelligence and achievement measures. The intelligence estimate used for correlation purposes was the full scale or global score on the particular instrument administered.

Of the cases analyzed, 73% were given the Wechsler Intelligence Scale for Children-Third Edition (Wechsler, 1991); 22% the Woodcock-Johnson Psycho-educational Battery - Revised (Woodcock & Johnson, 1989); and 5% the Stanford-Binet Intelligence Scale-Fourth Edition (Thorndike, Hagen, & Sattler, 1986).

The achievement estimates used for correlation purposes were percentile ranks converted to the same eight category scale used in the survey. The primary achievement batteries utilized included the Kaufman Test of Educational Achievement (Kaufman & Kaufman, 1985) (46%), Woodcock-

Johnson Psycho-educational Battery - Revised (Woodcock & Johnson, 1989)(39%), Test of Early Math Ability (Ginsburg & Baroody, 1983)(7.5%), and Test of Early Reading Ability (Reid, Hresko, & Hammill, 1981)(7.5%).

Due to partial data, correlation coefficients could not be derived for all areas. Nonetheless, the coefficients derived for intelligence and six areas of achievement are listed in Table 2.

Insert Table 2 about here

As the reader will note, all of the correlation coefficients were in the moderate to high range and statistically significant at the .05 level. These findings indicate that the teachers in this investigation were very accurate predictors of student performance on standardized tests. The .74 correlation found between teachers' point estimates of cognitive ability and students' full scale scores on intelligence tests is of particular interest in that it is stronger than the majority of those previously reported in the literature. Further highlighting the accuracy of estimates supplied by the teachers surveyed in this study is the fact that there was a mean difference of eight standard score points between teachers' point estimates and students' actual full-scale intelligence quotients and that the mean for teacher estimates (92.85)

was virtually identical to the mean for students' actual IQ scores (92.9).

Given that the correlation coefficients found between teacher estimates and actual test scores cluster around the mean correlation of .55 reported for studies of this kind (Hoge & Butcher, 1984), it would appear that the answer to our first research question is "yes," the teachers in this study were able to predict the test performance of student referred for special education at a level of accuracy that was comparable to or higher than the standards reported in the literature. Furthermore, the results of this investigation as they pertain to our second research question would support the conclusion that teachers can be highly accurate in estimating student performance based on the standard score and percentile rank formats commonly used to convey the results of psychometric tests.

Analysis of Variance/T-Test Results

Analysis of variance and t-test results were remarkably consistent across all areas of this investigations. Tables 3 and 4 contain a summary of these findings with respect to the seven dependent variables for which complete data were available (teacher estimates of intelligence, math computation, math overall, reading decoding, reading comprehension, reading overall. and spelling) and seven independent variables (student age,

grade, and gender; length of affiliation between teacher and student; and teachers' level of training and experience).

With regard to the accuracy of teachers' ability to predict student performance on standardized tests, differences were found in only two areas (Reading Decoding and Reading Comprehension) and with respect to one independent variable (teacher level of training). More specifically, teachers trained at the "MA level or higher" were found to be significantly more accurate than those at the Bachelors level in estimating student performance in these two specific domains ($p<.03$ for both variables). None of the other seven independent variables listed in Tables 3 and 4 were found to vary significantly with respect to the dependent variables.

Insert Tables 3 and 4 about here

Analysis of variance and t-test results also revealed no significant differences between those who had reviewed the results of achievement and intelligence tests prior to making their estimates. These findings suggest that exposure to educational test data did not affect the accuracy of teacher estimates and support the contention that the respondents' ratings were based primarily on professional judgment and experience with the student.

Hit Rates

Hit rates for teacher estimates of student performance on standardized intelligence tests were obtained by matching range estimates with actual scores expressed in ranges. Since the data were obtained from several instruments and at multiple age levels, the standard errors of measurement for full scale scores varied from three to six points. The following three general categories were, therefore, derived to account for these variations and establish a basis for comparison: Highly Accurate (defined as +/- 3 standard score points), Accurate (defined as +/- 4 to 12 standard score points), and Inaccurate (defined as +/- 13 or more standard score points). Results indicate that 18.2% of the teacher estimates obtained in this investigation fell in the "Highly Accurate" range while 61.3% and 20.5% fell in the "Accurate" and "Inaccurate" categories respectively. Generally speaking, these results indicate that four out of five teachers surveyed were able to provide an IQ estimate that fell within plus or minus 12 standard score points of a students actual performance on an intelligence test.

Discussion

The results of this study indicate that teachers can be extremely accurate predictors of the standardized test performance of students referred for special education. They, therefore, appear to have a clear understanding of

student strengths and weakness at the time of referral. This finding has important implications for referral and placement practices.

First of all, the fact that the results of this study were achieved with a group of students that would be described as "low achieving" and by using a rating procedure similar to that employed by school psychologists and educational diagnosticians indicates that teacher judgment can be a very important adjunct to psychometric testing in the special education diagnostic and placement process. This seems particularly true in light of the fact that the high degree of accuracy displayed by teachers did not vary significantly according to most teacher or student demographic parameters.

In particular, teacher estimates may provide an important source of verification for the results of psychometric instruments. A high rate of agreement between test scores and teacher estimates would strengthen an examiner's confidence in the results of an evaluations while a significant discrepancy would be grounds for questioning them. This is not to suggest that a marked discrepancy between teacher estimates and student performance on standardized assessments would necessarily discount the meaningfulness of test data. It would, however, indicate that the investigation of such a discrepancy should be an important clinical task that may lead to interesting and useful hypotheses about children's functioning. This

approach would be consistent with the conclusions of Gresham et al. (1987) who argued that teachers are an important source of validation for the results of psychometric testing.

Secondly, the results of this investigation support the conclusion of Sharpley and Edgar (1986) who reported that the low to moderate correlation coefficients between teacher rankings and performance on standardized tests indicate that while the two sources of information overlap dramatically, they also contribute their own unique and valuable information. Therefore, it seems reasonable for multi-disciplinary diagnostic teams to give teacher judgment an influential role in the identification/placement process. One possible format for this might be the use of standardized teacher questionnaires. Hoge (1983,) in a review of these instruments, found them to be psychometrically sound and useful as supplements to traditional psycho-educational batteries.

Thirdly, the results of this investigation affirm the roles of specialists (i.e., School Psychologists, Educational Diagnosticians, and School Social Workers) in the pre-referral stage of the special education eligibility process. While teachers appear to be quite accomplished at estimating the test performances of their students, prior research suggests that they may not be accurate judges of the existence of handicapping conditions (Gerber & Semmel, 1984; Rothlisberg & Liljestrom, 1984). Therefore, given the

importance of initial referral decisions to the integrity of diagnostic and placement practices (Ysseldyke & Algozzine, 1981; Ysseldyke, Algozzine, Regan, & McGue, 1981), it would seem essential that those knowledgeable in the application of eligibility criteria be involved at the very earliest stages of a referral..

This investigation has also found that teachers are somewhat ambivalent about the information they receive from psycho-educational reports. This conclusion was evidenced by the fact that a majority of respondents agreed with two items suggesting that intelligence and achievement testing contributed positively to their understanding of and efforts to teach referred students but gave much more varied responses to an item suggesting that the primary purpose of intelligence/achievement testing was to qualify students for special education. In addition, the lack of variation in mean scores on these items suggested that teachers were as likely to perceive psycho-educational assessment as "institutionally necessary" (that is, required in order to determine eligibility for special services) as they were to find it "instructionally useful" (that is, a source of information that assists in determining how best to teach a child).

In applying these findings to the special education classification and placement process, it appears that school psychologists and other diagnosticians would do well to encourage greater input from classroom teachers and refine

the prescriptive content in their reports. Psycho-educational reports that focus primarily or exclusively on the results of standardized tests do not offer much "new" information to teachers on the students they refer and fail to address the question that is implicit in each special education referral; that is, "How do I most effectively teach this child?" This conclusion is consistent with the literature in this area which suggests that teachers prefer detailed descriptions of students strengths and weaknesses and specific recommendations written in educationally meaningful terms (Weiner, 1985, 1987).

Furthermore, since the results of this study indicate that teachers know a great deal about a student's cognitive and academic functioning at the time of referral, it would seem that the most effective evaluation procedure would be one with an in-depth focus on areas that teachers may not be as adept at evaluating (e.g., information processing and personal-social development). The results of a study by Argulewicz and Miller (1985) would support such an approach. They found that teachers' ability to estimate student functioning in the psycho-emotional domain, as demonstrated by their ability to predict student performance on a standardized anxiety scale, was extremely low. Therefore, broad focused evaluations that describe a wider range of behavior than just intelligence and achievement would truly add to a teacher's knowledge of the student and better address the ultimate purpose of a special education

referral; that is, the development of a meaningful plan of action.

Finally, the results of this investigation have implications for teacher empowerment in the implementation of special education reform. Meijer and Foster (1988) and Gutkin and Hickman (1988) found that teacher control and self efficacy had a significant impact on their openness to consultation. Given that consultation is crucial to the integration of students with disabilities into mainstream classrooms, it would seem beneficial for school psychologists to encourage teachers to perceive themselves as efficacious in solving typical student problems. One way to accomplish this would be to increase their sense of control over student problems by empowering them in the decision-making process with at-risk students. Giving teachers the recognition, status, and influence of first-rate diagnosticians would be an important contribution to this end.

To complete this discussion, several methodological limitations to this study must be acknowledged. First of all, it focused on a relatively small, geographically limited sample. Readers are, therefore, cautioned against over-generalizing its findings. Secondly, because the sample of teachers that responded to the survey was highly trained and experienced, they may not be typical of teachers in general. Therefore, while only minimal differences were found between the accuracy of estimation of less

experienced, bachelors level teachers and those with more experience and advanced degrees, these results are best interpreted as indicating the predictive efficacy and attitudes of experienced, masters-level educators. Lastly, the sample of teachers who responded to this survey had done a considerable amount of pre-referral work with the students whose standardized test performance they were estimating. This experience may have given them an unusually strong basis for their predictions.

Conclusions

It is critical that psychological and educational evaluations completed on students during the referral process give careful consideration to the knowledge that teachers already have as well as provide information that is instructionally useful. The data from this investigation indicate, first and foremost, that teachers can accurately predict their students' performance on standardized tests of intelligence and achievement using the same standard score and percentile rank formats employed by school psychologists and other diagnosticians. In addition, the results of an attitudinal questionnaire administered in this investigation are consistent with the literature on teacher perceptions of psycho-educational evaluations in suggesting that teachers are as likely to view intelligence and achievement testing as "institutionally necessary" as they are to see them as "instructionally useful." Therefore, it would appear that

the traditional emphasis in psychological and educational reports on intelligence and achievement test scores, rather than practical interventions based on those scores, may be inappropriate. To improve referral and placement practices in special education, the investigators would, as a result, argue for greater teacher input into eligibility decisions, more in-depth involvement of specialists in the pre-referral stage, and added emphasis on assessment domains that are more likely to yield new and useful information for teachers. These include such areas as affective development, behavioral adjustment, motivation, and data-based treatment recommendations. These changes would be consistent with the goals of promoting the integrity of identification and placement procedures, empowering teachers, and improving the instructional relevance of psycho-educational evaluations.

References

Argulewicz, E. & Miller, D. (1985). Self-reported and teachers' rankings of anxiety among first-grade children. School Psychology Review, 14, 75-78.

Coladarci, T. (1986). Accuracy of teacher judgments of student responses to standardized test items. Journal of Educational Psychology, 78, 141-146.

Comprehensive Test of Basic Skills (1984). Monterrey, CA: CTB MacMillan/McGraw-Hill

Comprehensive Test of Basic Skills (1990). Monterrey, CA: CTB MacMillan/McGraw-Hill

Dusek, J.B. & O'Connell, E.J. (1973). Teacher expectancy effects on the performance of elementary school children. Journal of Educational Psychology, 65, 371-377.

Gerber, M.M. & Semmel, M.I. (1984). Teacher as imperfect test: Reconceptualizing the referral process. Educational Psychologist, 19, 137-148.

Gresham, F., Reschly, D., & Carey, M. (1987). Teachers as "tests": Classification accuracy and concurrent validation in the identification of learning disabled children. School Psychology Review, 16, 543-553.

Gutkin, T. & Hickman, J. (1988). Teachers' perceptions of control over presenting problems and resulting references for consultation versus referral services. Journal of School Psychology, 26, 395-398.

Hoge, R. (1983). Psychometric properties of teacher judgment measures of pupil attitudes, classroom behaviors, and achievement levels. Journal of Special Education, 17, 401-429.

Hoge, R. & Butcher, R. (1984). Analysis of teacher judgments of pupil achievement levels. Journal of Educational Psychology, 76, 777-781.

Hopkins, K., George, C., & Williams, D. (1985). The concurrent validity of standardized achievement tests by content area using teachers' ratings as criteria. Journal of Educational Measurement, 22, 177-182.

Jatak, S.R., & Wilkinson, G.S. (1985). The Wide Range Achievement Test-Revised. Wilmington, DE: Guidance Associates.

Kaufman, A.S., & Kaufman, N.L. (1985). The Kaufman Assessment Battery for Children. Circle Pines, MN: American guidance Service.

Kellaghan, T., Madaus, G., & Airasion, P. (1982). The effects of standardized testing. Boston: Kluwer-Nijhoff Publishing.

Luce, , S.R., & Hoge, R.D. (1978). Relations among teacher rankings, pupil-teacher interactions, and academic achievement: A test of the teacher expectancy hypothesis. American Educational Research Journal, 15, 489-500.

Mayfield, B. (1979). Teacher perception of creativity, intelligence, and achievement. Gifted Child Quarterly, 23, 812-817.

Meijer, C. & Foster, S. (1988). The effect of teacher self-efficacy on referral chance. The Journal of Special Education, 22, 378-385.

Morine-Dershimer, G. (1978-79). The anatomy of teacher prediction. Educational Research Quarterly, 3, 59-65.

Mulholland, L.A. & Berliner, D.C. (1992). Teacher experience and the estimation of student achievement. ERIC Document Reproduction Service No. ED 348 350)

Noble, B. & Dickinson, D. (1988). Utility of school psychologists' recommendations: Perception of elementary teachers. Psychology in the Schools, 25, 412-418.

Oliver, J.E., & Arnold, R.D. (1978). Comparing a standardized test, an informal inventory, and teacher judgment of third grade reading. Reading Improvement, 15, 56-59.

Ownby, R., Wallbrown, F., Brown, D. (1982). Special education teachers' perceptions of reports written by school psychologists. Perceptual and Motor Skills, 55, 955-961.

Prescott, G.A., Barlow, I.H., Hogan, T.P., & Farr, R.C. (1987). The Metropolitan Achievement Test (6th ed.). San Antonio, TX: The Psychological Corporation.

Quay, L.C. (1996). The developmental rating scale: A teacher-friendly alternative to commercial assessment devices. (ERIC Document Reproduction Service No. ED 397 114)

Reid, D.K., Hresko, W.P., & Hammill, D.D. (1981). Test of Early reading Ability. Austin, TX: PRO-ED.

Reilly, T., Drudge, O., Rosen, J., Loew, D., & Fischer, M. (1985). Concurrent and predictive validity of the WISC-R, McCarthy Scales, Woodcock-Johnson, and academic achievement. Psychology in the Schools, 22, 380-382.

Rothlisberg, B. & Liljestrom, M. (1984). Recognizing the learning disabled child: Which behaviors do teachers use? Paper presented at the annual meeting of the American Psychological Association, Toronto, CAN. (Also in ERIC Document Reproduction Service No. ED 250 839)

Schendel, J. & Binder-Reschly, M. (1989). Criterion validity of curriculum based assessment and correlation with teacher ratings and ITBS scores. A research project of the Iowa learning disabilities research consortium. (ERIC Document Reproduction Service No. ED 318 152)

Semmel, M.I. (1987). Special education in the year 2000 and beyond: A proposed action agenda for addressing selected ideas. The Future of Special Education: Proceedings of the Council for Exceptional Children, Lake Geneva, WI. (Also in ERIC Document Reproduction Service No. ED 192 110).

Sharpley, C. & Edgar, E. (1986). Teachers' ratings vs standardized tests: An empirical investigation of agreement between two indices of achievement. Psychology in the Schools, 23, 106-111.

Stader, E. (1990). Expert and novice teachers' ability to judge student understanding. (ERIC Document Reproduction Service No. ED 322 127)

Stevenson, H., Parker, T., Wilkinson, A., Hegion, A., & Fish, E. (1976). Predictive value of teachers' ratings of young children. Journal of Educational Psychology, 68, 507-51

Thorndike, R.L., Hagen, E.P., & Sattler, J.M. (1986). The Stanford Binet Intelligence Scale-Fourth Edition. Riverside, CA: The Riverside Pub. Co.

Thurlow, M.L., Christenson, S., & Ysseldyke, J.E. (1983). Referral research: An integrative summary of findings (Research Report No. 141). Minneapolis: University of Minneapolis.

Valus, A. (1986). Teacher perceptions of identification criteria emphasized in initial learning disabilities placement. Learning Disabilities Research, 2, 21-25.

Webster, J., Hewett, B., & Crumbacker, M. (1989). Criterion-Related validity of the WRAT-R and K-TEA with teacher estimates of actual classroom academic performance. Psychology in the Schools, 26, 243-248.

Wechsler, D. (1974). Wechsler Intelligence Scale for Children - Revised. San Antonio, TX: Psychological Corporation.

Wechsler, D. (1991). Wechsler Intelligence Scale for Children - Third Edition. San Antonio, TX: Psychological Corporation.

Weiner, J. (1985). Teachers comprehension of psychological reports. Psychology in the Schools, 22, 60-64.

Weiner, J. (1987). Factors affecting educators' comprehension of psychological reports. Psychology in the Schools, 24, 116-126.

Wilson, J., & Wright, C.R. (1993). The predictive validity of student self-evaluations, teacher assessments, and grades for performance on the verbal reasoning and numerical ability scales of the differential aptitude test for a sample of secondary school students attending rural appalachia schools. Educational and Psychological Measurement, 53, 259-70.

Woodcock, R.W., & Johnson, M.B. (1977). Woodcock-Johnson Psycho-Educational Battery. Allen, TX: DLM Teaching Resources.

Woodcock, R.W., & Johnson, M.B. (1989). Woodcock-Johnson Psycho-Educational Battery -Revised. Allen, TX: DLM Teaching Resources.

Ysseldyke, J., Algozzine, B., Regan, R., & McGue, M. (1981). The influence of test scores and naturally occurring pupil characteristics on psycho-educational decision making with children. Journal of School Psychology, 19, 167-177.

Ysseldyke, J. & Algozzine, B. (1981). Diagnostic classification decisions as a function of referral information. The Journal of Special Education, 15, 429-435.

Ysseldyke, J. & Shinn, M.R. (1981). Psycho-educational evaluation. In J.M. Kauffman & D.P. Hullahan (eds.), Handbook of special education (pp. 418-440). Englewood Cliffs, NJ: Prentice-Hall.

Zins, J.E. & Barnett, D.W. (1983). Report writing: Legislative, ethical, and professional challenges. Journal of School Psychology, 21, 219-227.

Table 1: Descriptive Statistics for Attitudinal Items

Attitudinal Items:	Mean	Median	Standard Deviation
1. Information from intelligence and achievement testing improves my understanding of students.	6.3	7.0	1.8
2. Information from intelligence and achievementtesting helps me to teach more effectively.	6.0	6.0	2.0
3. The primary purpose of intelligence and achievement testing is to determine student eligibility for special education programs.	5.3	6.0	2.0

Table 2: Correlations Bewteen Teacher Estimates and Actual Test Scores

		ACTUAL SCORES							
		IQ Point	IQ Range	Reading Comprehension	Reading Decoding	Overall Reading	Math Computation	Overall Math	Spelling
	IQ Point	.74							
	IQ Range		.56						
TEACHER	Reading Comprehension			.53					
ESTIMATES	Reading Decoding				.42				
	Reading Overall					.74			
	Math Computation						.41		
	Math Overall							.47	
	Spelling								.50

BEST COPY AVAILABLE

Table 3: Summary of Analysis of Variance Results

Independent Variables	Dependent Variables	F	F Prob.
Student Age	Estimates of:		
Under 7.5 years	Intelligence	.1125	.8939
7.51 - 10 years	Math Computation	2.5811	.0957
10+ years	Math Overall	.2038	.8172
	Reading Decoding	.9098	.4146
	Reading Comprehension	.1800	.8363
	Reading Overall	.1350	.8744
	Spelling	1.5774	.2290
Length of Affiliation			
0 - 3 months	Intelligence	.0518	.9496
4 - 6 months	Math Computation	2.1448	.1366
7+ months	Math Overall	.0776	.9255
	Reading Decoding	.5993	.5559
	Reading Comprehension	.6072	.5516
	Reading Overall	1.6765	.2082
	Spelling	.1252	.8829
Student Grade			
Pre-K to 2	Intelligence	.3244	.7248
3 to 5	Math Computation	2.8712	.0740
6 or higher	Math Overall	.8948	.4219
	Reading Decoding	.0064	.9937
	Reading Comprehension	.3682	.6952
	Reading Overall	.1598	.8533
	Spelling	1.2011	.3190

Table 4: Summary of T-Tests Results

Independent Variables	Dependent Variables	T	2-Tail Prob.
Student Gender	Estimates of:		
Male	Intelligence	1.92	.301
Female	Math Computation	1.46	.670
	Math Overall	2.94	.087
	Reading Decoding	2.33	.295
	Reading Comprehension	3.21	.117
	Reading Overall	1.15	.958
	Spelling	2.56	.141
Teacher Training			
Less than MA	Intelligence	1.36	.588
MA or higher	Math Computation	1.08	.838
	Math Overall	1.23	.666
	Reading Decoding	3.13	.030
	Reading Comprehension	4.24	.030
	Reading Overall	1.20	.715
	Spelling	1.48	.560
Teacher Experience			
0- 10 years	Intelligence	1.17	.684
10+ years	Math Computation	1.64	.615
	Math Overall	1.65	.678
	Reading Decoding	2.69	.077
	Reading Comprehension	1.54	.672
	Reading Overall	1.84	.518
	Spelling	1.49	.473